The Autumn Experiment 2013: a Pilot Study of a Forecasting Test Bed at the Met Office

Forecasting Research
Technical Report No: 596

November 2014
Adrian Semple

The Autumn Experiment 2013: A Pilot Study of a Forecasting Test Bed at the Met Office

Abstract

The Autumn Experiment 2013 explored the feasibility and suitability of the test bed environment for Science and Forecasting at the Met Office. It was the first initiative of its kind at the Met Office and ran for two weeks from the 18th to the 29th November 2013. Its overriding aim was as a pilot study in providing an environment in which Operational Meteorologists could work alongside Research Scientists and evaluate the Met Office's forecasting capability. All aspects of the forecasting process were open to evaluation: the numerical model output and its display and interpretation, the use and interpretation of satellite imagery, forecasting techniques and judgement calls, the messaging of weather risks to the public and the associated issuing of weather warnings.

This paper documents and evaluates the experimental procedure developed in this test bed and makes recommendations so that the initiative can be developed further in future. The report goes on to summarise the findings of the experiment with a focus on the model problems identified within the higher level model skill normally observed so that these may be addressed by relevant research or post-processing groups. The report also documents in more detail three forecast cases from the experiment which exemplify the sort of issues which can be identified in this environment and also go to illustrate the benefits that a detailed knowledge of the numerical models are to the forecasting process.

Finally, the experiment is summarised in terms of its outcomes and benefits to the Met Office in terms of improving its forecasting capability through the increased interaction between Operational Meteorologists and Research Scientists in the test bed environment.

Contents

The Autumn Experiment 2013: A Pilot Study of a Forecasting Test Bed at the Met Office 2
1. Introduction 5
2. Experiment Objectives 6
3. Experiment Design and Process 6
Methodology 6
Personnel 7
Room Layout 9
Schedule 9
Models Available for Evaluation 12
Ensembles Available for Evaluation 14
4. Forecast Routine 14
5. Highlight Forecast Issues 15
Over- forecasting of widespread dense fog 15
Forecast Issued 25th November 2013, Forecast Verified 26th November 2013 15
Verification: 19
T+12 to T+18 19
T+18 to T+30 19
Corrupted Visibilities by Post Processing 22
Forecast Issued 28th November 2013 22
Forecast temperatures too low under ridge conditions and clear skies – decoupling in light wind regimes 24
Forecast issued 22nd November 2013 24
6. Summary of Findings 27
Evaluation of Existing and Experimental Met Office Modelling Capability 27
UKV Op v UKV BLayer Observations 27
UKV Op v Euro4 Observations 27
Known Model Characteristics Observed 28
Notable Poor Forecasts 28
Identified Process Changes 28
Suggested Product Improvements 29
GPP 29
Simulated Satellite Imagery 29
Clive Wilson's Model Comparisons 29
Forecast Monitoring 29
Ensembles 30
Suggested New Product Requirements 30
Noted Product inefficiencies/limitations 31
Noted Products which Decreased Confidence 31
7. Outcomes 31
Process 31
Models & Products 31
Knowledge & Development 32
8. Benefits 32
9. Recommendations 34
10. Lessons Learned 35
11. Summary 37
12. Acknowledgements 39

13. References....39
14. Appendix....40
Forecast and Evaluation Templates....40
1. Latest Model critique against current observations (00z-12z)....41
2. Forecast Guidance & Evidence....41
3. Model/Product/Technique Evaluation....42
1. Model critique against current conditions....44
2. Forecast Guidance & Evidence....44
3. Model/Product/Technique Evaluation....45
15. Testimonials....49

1. Introduction

The Autumn Experiment took place for a two week period between 18-29 November 2013 and involved 18 participants from across Science, Forecasting, Public Weather Service (PWS) and Training. The experiment's primary aim was that of a feasibility study in conducting an exercise bringing Scientists and Operational Meteorologists together so as to assess the Met Office's forecasting capability. All aspects of the forecasting process were open to evaluation; observations, models, products, forecasting techniques and warnings.

The Autumn Experiment was inspired by the emergence of "weather test beds" in other National Met Services (particularly in the United States and Europe) which have largely developed in the last ten years (e.g. Ralph. F.M *et al*. (2013)). Many of these experiments have grown out of a requirement to focus effort on the modelling capability of high-impact weather; e.g. hurricane prediction, severe storms, and extreme precipitation whilst others have developed to focus on other forecasting challenges ranging from mesoscale modelling to climate forecasting.

In essence the primary function of the test bed is to create a unique environment in which forecasting capability can be critically assessed. The test bed environment is unique because it simulates the operational environment and in some sense some of the pressures and constraints that operational meteorologists face but does so in a way which allows more time for critical discussion and an evaluation of the models, products and techniques used in the generation of the forecast than would otherwise be afforded in a true operational environment. Importantly it also brings together two groups of people which usually operate surprisingly independently: operational meteorologists and research scientists. By bringing research scientists into the simulated operational environment they have the opportunity to experience the challenges of operational forecasting and take part in forecasting using models and products which they have been involved in developing. The result is an exchange of ideas and insights which are mutually beneficial to both forecasting and numerical weather prediction.

As discussed above, most weather test beds focus on one particular aspect of meteorology; a particular challenging and often high-impact weather phenomenon that has an identified gap in the existing forecasting or scientific capability. The chosen phenomenon will naturally be either critical to forecasting success or else have a high impact in the country in which the test bed is taking place.

When the Autumn Experiment was devised it was primarily seen as a pilot study from which the Met Office would develop the initiative in future years. The focus of the experiment was therefore as much about the experimental process as it was on the forecasting capability. Further, since the weather in the UK is hugely variable for any given time of year and lacks a dependable

weather aspect on which to focus on, it was decided to open up the experiment to focus on the forecasting challenges from T+0 to T+144 and to use the relevant Met Office models for the appropriate time ranges. To this end the day was divided into two periods: the 'short-period' from T+0 to T+48 and the 'medium-range' from T+60 to T+144. This provided scope to gain experience in assessing models of different resolutions and also ensembles at different forecast ranges. It also provided the opportunity to assess a wealth of Met Office products which are available throughout the forecast ranges.

2. Experiment Objectives

As discussed earlier, the primary goal of the experiment was to perform a feasibility study on whether the experimental environment would be effective for the Met Office. To be shown to be effective, however, the experiment must also show signs of accomplishing several other objectives which form the basis of the model and product assessment and evaluation activities. The experiment objectives were thus:

- To establish an experimental process in which to evaluate the forecasting process for this and future experiments.
- To evaluate and develop new and current scientific models, techniques and products.
- To evaluate and develop forecasting and warning techniques using a variety of numerical model output from both deterministic and ensemble forecasting systems.
- To increase pull-through of science products and techniques to operations.
- To identify new forecaster requirements.
- To help guide forecasting and science strategy.

The last of these objectives is perhaps one of the most important because it illustrates how the effectiveness of the experimental method extends from the refinement of existing models and products to the long-term strategic direction of the Met Office.

3. Experiment Design and Process

Methodology

The experimental process consisted of two groups of eight people with mixed disciplines performing parallel forecasting and evaluation activities. Parallel activities were ran both to achieve optimum group size (thereby promoting open discussion) and also to allow for the possibility that differences would appear in the two forecasts that could be identified and discussed. The members of each group were changed at the start of each day so as to allow people with different skill-sets to work together. This method of parallel forecasting and evaluation activities proved to be very successful with open

discussion among the group members producing a unique environment in which knowledge and experience could be shared effectively and productively.

Personnel

The experimental personnel consisted of six operational meteorologists, seven research scientists (of which one was the project lead), one Public Weather Service advisor, and one Met Office College trainer. Of the operational meteorologists, two were experienced Chief Forecasters. Although most of the experiment participants attended the experiment throughout the two week period, there were some differences to accommodate personal circumstances. The experimental personnel for the two weeks is listed below.

Week 1

Project Lead (1)	Forecast Monitoring and Investigations	Adrian Semple
Forecasting (6)	Chief Forecaster	Steve Willington
	Chief/Deputy Chief Forecaster	Bob Wilderspin
	Forecaster (Team Leader)	Mark Jellis
	Forecaster	Matthew Lewis
	Forecaster	Katie O'Boyle
	Forecaster (BBC)	Laura Gilchrist
Science (6)	Numerical Modeling: Mesoscale Model Development	F. Jorge Bornemann
	Numerical Modeling: Model Diagnsotics and Novel Verification	Rachel North
	Numerical Modeling: Weather Impacts	Rob Neal
	Numerical Modeling: Weather Impacts	Helen Titley
	Weather Science: Data Assimilation & Ensembles	Chiara Piccolo
	Satellite Applications: Satellite Imagery Applications	Thomas Blackmore
PWS (1)	Civil Contingency Advisor (Scotland and NI)	Catherine Sweetman
College (1)	FPD Manager - Satellite and NWP Training	Deborah Miles

Table 1 Personnel for week 1 of the Autumn Experiment.

Week 2

During the second week of the experiment a guest attendee was included. This was from a response to an invitation to people, generally management or senior management, who would like to be involved but could not dedicate a full week to the experiment. This had the advantage that managers or senior researchers could experience first hand what the experiment was about.

Project Lead (1)	Forecast Monitoring and Investigations	Adrian Semple
Forecasting (6)	Chief Forecaster	Steve Willington
	Chief/Deputy Chief Forecaster	Bob Wilderspin
	Forecaster (Team Leader)	Mark Jellis
	Forecaster	Matthew Lewis
	Forecaster	Katie O'Boyle
	Forecaster (BBC)	Laura Gilchrist
Science (6)	Numerical Modeling: Mesoscale Model Development	F. Jorge Bornemann
	Numerical Modeling: Model Diagnsotics and Novel Verification	Rachel North
	Numerical Modeling: Weather Impacts	Rob Neal
	Numerical Modeling: Weather Impacts	Helen Titley
	Data Assimilation & Ensembles: Earth System DA	Brett Candy
	Satellite Applications: Satellite Imagery Applications	Thomas Blackmore
PWS (1)	Civil Contingency Advisor (S, W and Wales)	Malcolm Weatherall
College (1)	FPD Manager - Satellite and NWP Training	Deborah Miles
Guest (1)	Guest attendee – one visitor per morning or afternoon session.	Keith Williams Bruce Wright David Walters Huw Lewis Rutger Dankers Ken Mylne Jon Taylor

Table 2 Personnel for week 2 of the Autumn Experiment.

Room Layout

The room layout during the experiment is shown in Figure 1. Met Office college training rooms 1 and 2 were both used for the experiment but with the room divider removed. The majority of the workstations were also removed except for a main PC at the end of either room which was connected to the overhead projector and two PCs nearby. The PC with the overhead projector was used as the main means to perform the meteorological analysis and provided an effective way for the group to discuss the issues. The two nearby PCs were used to partially investigate side issues which arose from the group discussion, but also as the main means of recording the group's findings in the experiment templates which were provided.

The intention of the room layout was to provide two distinct areas for two groups to operate in whilst still maintaining the unity of the experimental group as a whole. This also allowed for groups to come together at the end of the activities for open discussion about the particular challenges of the day.

Schedule

The following schedule is a revised version that was used by the end of the experiment. One of the aims of the experiment was to establish a process that could be used for further forecasting experiments in the future, and one of the core things to investigate was the experimental routine.

The day was split into two halves - the morning being focussed on the short-range period to T+48 and the afternoon on the medium-range to T+144. The idea was that at the end of each forecasting period the groups would verify and discuss the previous day's forecast. In the event it was found that producing the forecast from T+0 to T+144 actually took more time than was envisaged because of the sheer wealth of models and model products which were available and because of the amount of discussion taking place. Although valuable, this meant that the verification activity was reduced to a duration that made the activity less useful than it could have been. To achieve the correct balance in the next experiment, it is envisaged that the forecasting scope will be significantly reduced – this is one of the main lessons learned from this pilot experiment.

The Chief's Brief was introduced at the start of the day by the end of the first week. This provided an initial gathering of the experiment participants each day and enabled initial areas of interest to be identified as a group.

The medium-range forecasts naturally had a different set of challenges to the short-range forecasts; the scope of this period alone is enormous given the degrees of freedom in the solutions at the longer ranges. For this reason, it was largely found that there was barely enough time to consider the medium range in the afternoons of the experiment. Verification of the medium range period was also difficult given the time constraints and the fact that each

forecast didn't verify for days after the forecast was made. For these reasons the medium-range forecasting period was limited during this pilot study.

In all cases, the latest runs of the relevant models were used for the evaluation exercises.

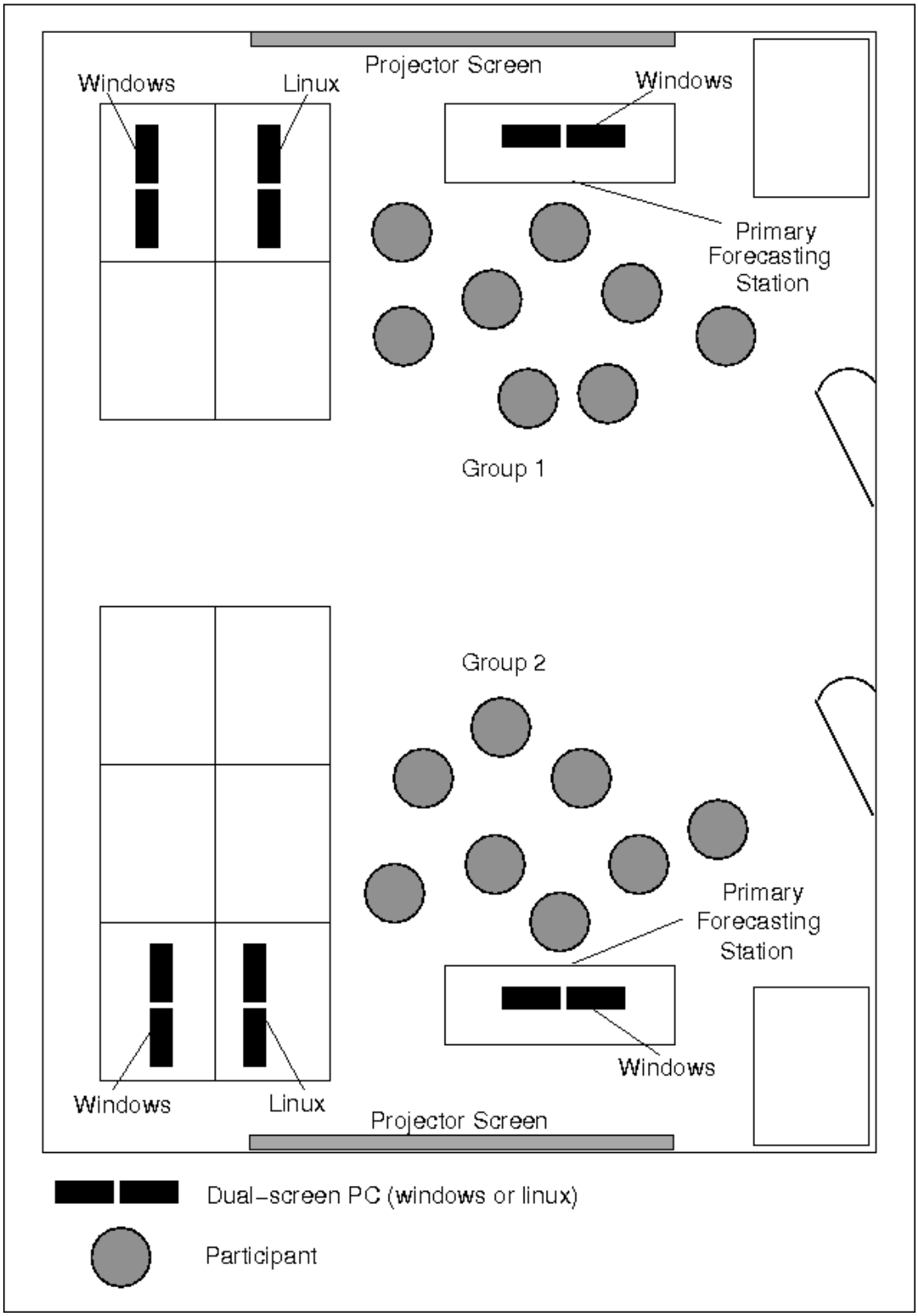


Figure 1 Room layout for the Autumn Experiment 2013

09:00-09:20	**Chief's Brief** Chief Forecaster gives an overview of T+0 to T+144, identifying any particular areas of interest**.**
09:20-11:10	**Production of UK forecast guidance from T+12 to T+48 and evaluation of models/products/techniques used.** Available: i) Deterministic models: QG00 GM, QV06 UKV, E400 Euro4, London Model ii) Ensembles: EG00 MOGREPS-G, EK03 MOGREPS-UK, ECMWF, multi-model ensemble iii) Products: Difference-Tracking, simulated satellite imagery, EPS-W, Regime Analysis iv) Observations: satellite imagery, observations. Issue text Forecast Guidance (Word template) including decisions made illustrated with model fields as evidence. Includes: model critique for the latest runs at the start of the period, discussion of messaging and warnings, and completion of the warnings matrix.
11:10-11:20	**Forecast Discussion** Open discussion between groups of main decision challenges, model guidance and decisions made
11:20-11:40	**Break**
11:40-12:40	**Evaluation of yesterday's Period 1 Forecast (T+12 to T+48)** Yesterday's forecast groups subjectively evaluate their forecast and model guidance. Include fields from yesterday with the verifying fields. Comment on the performance of the models. *In light of verification and with the benefit of hindsight, note any evidence which could have been interpreted differently. Could this be used differently in future?*
12:40-13:30	**Lunch**
13:30-14:00	**Evaluation Discussion** Each group discusses significant points of interest about the verification of yesterday's forecasts.

Table 3 Schedule for the morning session.

14:00-15:15	**Production of UK forecast guidance from T+48 to T+144 and evaluation of models/products/techniques used.** Available: i) Deterministic models: QG00 GM, E400 Euro4 ii) Ensembles: EG00 MOGREPS-G (to T+72), MOGREPS-15 (to D+15), EK03 MOGREPS-UK, ECMWF, multi-model ensemble iii) Products: Difference-Tracking, Regime Analysis, EPS-W iv) Observations: satellite imagery, observations Issue text Forecast Guidance (Word template) including decisions made illustrated with model fields as evidence. Issue Warnings chart (Visual Weather) if appropriate.
15:15-15:30	**Break**
15:30-17:00	**Production of UK forecast guidance from T+48 to T+144 and evaluation of models/products/techniques used (CONTINUED).**

Table 4 Schedule for the afternoon session.

Models Available for Evaluation

The models available for evaluation during the experiment are shown in Table 5 and Table 6. The tables show important characteristics at the time of the experiment for comparison.

Additionally to those shown in the table, a version of the UKV with an experimental boundary layer scheme was available; this is referred to elsewhere in the text as ‘UKV Blayer’.

	Global Model	Euro4	UKV	London Model
Operational Status	Operational	Operational	Operational	Non-Operational
Horizontal Resolution	25Km (N512)	4.4Km	VR stretching (~1.5Km to ~4Km)	~330m
Vertical Levels	70L (80Km top)	70L (40Km top)	70L (40Km top)	70L
Run Times	00,06,12,18	00,06,12,18	00,03,06,09,12,15,18,21	03
Forecast Range	T+168 (0,12), T+67 (6,18)	T+120 (00,12), T+60 (06,18)	T+36	T+36
Timestep	10min	100secs	50secs	10secs
Land Surface Scheme	Jules	Jules	Jules	
Ancillaries	OSTIA (SST, Seaice), Ozone, Aerosols, Leaf area index	N/A	OSTIA (SST, Seaice), Ozone	N/A
DA Type	Hybrid Incremental 4D-Var	N/A (GM Downscaler - Reconfiguration from GM analysis)	Incremental 3D-Var + LHN nudging for precipitation	N/A (UKV Downscaler - Reconfiguration from UKV analysis)
DA Resolution	N216L70	N/A	~3Km	N/A
DA Time Window	6hrs	N/A	2hrs(IAU), 5hrs(AC)	N/A
Observations Assimilated	Sondes, wind profilers	N/A	Sondes, wind profilers	N/A
	Surface (stations, ships, buoys),		Surface (stations, ships, buoys),	
	Aircraft,		Aircraft,	
	ATOVS, SSMIS, AIRS, IASI, Geostationary radiances,		SEVIRI clear-sky radiances	
	AMV, Scatterometer / passive microwave radiometer surface winds,		AMV, Scatterometer / passive microwave radiometer surface winds,	
	GPSRO,		RADAR Doppler wind,	
	Ground-based GPS, WV		Ground-based GPS	
			MOPS (Cloud from UKPP, radar surface rainrate)	

Table 5 Models available for evaluation in the Experiment.

Ensembles Available for Evaluation

	MOGREPS-UK	MOGREPS-G	MOGREPS-15	ECMWF
Operational Status	Operational	Operational	Operational	Operational
Horizontal Resolution	2.2Km	30Km (N400 to T+72)	60Km	30Km (T+0 to T+240)/60km (T+240 to T+360)
Vertical Levels	70L	70L	85L	62L
Number of Members	12 (MOGREPS-G Downscaler)	24 (2*12 lagged last 2 runs)	24 (2*12 lagged last 2 runs)	51
Run Times	03,09,15,21	00,06,12,18	00,12	00,12
Forecast Range	T+36	T+66 (0,6,12,18)	T+360	T+360

Table 6 Ensemble Models available for evaluation in the Experiment.

4. Forecast Routine

The following routine was established by the end of the experiment as an efficient framework for evaluating available models and issuing forecast guidance. It is a synthesis made by the author as a result of carrying out the forecast evaluations on each day in the different forecasting groups.

GLOBAL MODEL OVERVIEW

- GLOBAL MODEL. Perform an assessment of the larger scale synoptic pattern using Global Model diagnostics.

- MODEL DIFFERENCES. Evaluate differences between Global Models from other centres to establish alignment/confidence of the Met Office Global Model.

- ENSEMBLES. Consider Global Model ensemble products for predictability impacts.

HIGH RESOLUTION REFINEMENT

- HIGH RESOLUTION. Evaluate higher resolution models for the shorter-range period and compare with Global Model.

- ASSESS VERIFIED PERIOD. Compare verified frames of latest runs of available models with observations (observations, satellite, radar) to assess reliability of each model and establish which model is likely to

provide the best guidance. Check tephigrams: compare model and actual. Check simulated satellite imagery.

- ENSEMBLES. Consider high resolution ensembles products for predictability impacts.

MODEL CHARACTERISTIC REFINEMENT

- MODEL CHARACTERISTICS. Consider model biases/characteristics and adjust forecast accordingly. (Use appropriate diagnostics if available (e.g. GPP rain/snow mix warm/cold bias scenario diagnostics) to assess impact of model biases.

5. Highlight Forecast Issues

In general the Met Office models provided good guidance during the two weeks of the experiment. The following forecast issues have been selected from the evaluation forms filled out during the experiment because they are good examples of one or more of the following:

- Model forecast bust.
- Problem caused by post processing.
- Good added value/judgement calls by the forecasting teams.

Over- forecasting of widespread dense fog

Forecast Issued 25th November 2013, Forecast Verified 26th November 2013

Poor UKV Op cloud distribution across the south of England during the day, followed by ***all models incorrectly developing dense freezing fog*** *patches across the Midlands during the night. All models struggled to represent correct cloud fields later in the short-period having large impacts on night-time temperatures. Forecasting teams had insufficient confidence to issue a fog warning despite clear model signal and this proved to be an excellent judgement call.*

Forecast Reasoning (Consolidated from Groups 1 and 2):

General anticyclonic conditions prevailing across the UK with a ridge extending north-westwards from the North Atlantic. Global Model vorticity fields (not shown) show a short-wave upper trough running southwards across the UK during the early period. GM tephigrams at T+9 (09Z 25th November) and T+24 (00Z 26th November) suggest that as the upper trough runs across the north of the UK, the boundary layer will lift for a short time with convective tops to -3C; this indicating a small chance of moderate

showers developing. In contrast, GM tephigrams across the south of the UK (Figure 2) indicate that the upper trough has little significant connection to the surface, during which time the boundary layer depth actually decreases.

The UKV Op and UKV BLayer T+6 forecast fields from 03Z 25th November 2013 (Figure 3) were verified against the 09Z-1030Z satellite imagery (Figure 4) and surface observations in order to determine which of the models were currently showing better modelling of the situation. On this occasion the UKV Op was seriously deficient in cloud across central and southern England whilst the UKV BLayer showed better cloud distribution. The UKV BLayer model was also supported by the Euro4.

Model tehphigrams were used to show that a temperature of ~10C was required to break up the Sc through convective over-turning. This was unlikely to be achieved readily with temperatures of 4C at 10am. The cloud across the south of the UK was therefore included in the guidance (against UKV Op).

Tephigrams indicated the boundary layer top to be 5-6000ft. This meant that cloud pushing in from the north-west would be unimpeded by high-ground in the west of the UK. Advection techniques were used to estimate the subsequent cloud distribution and the conclusions were consistent with the UKV Blayer forecast.

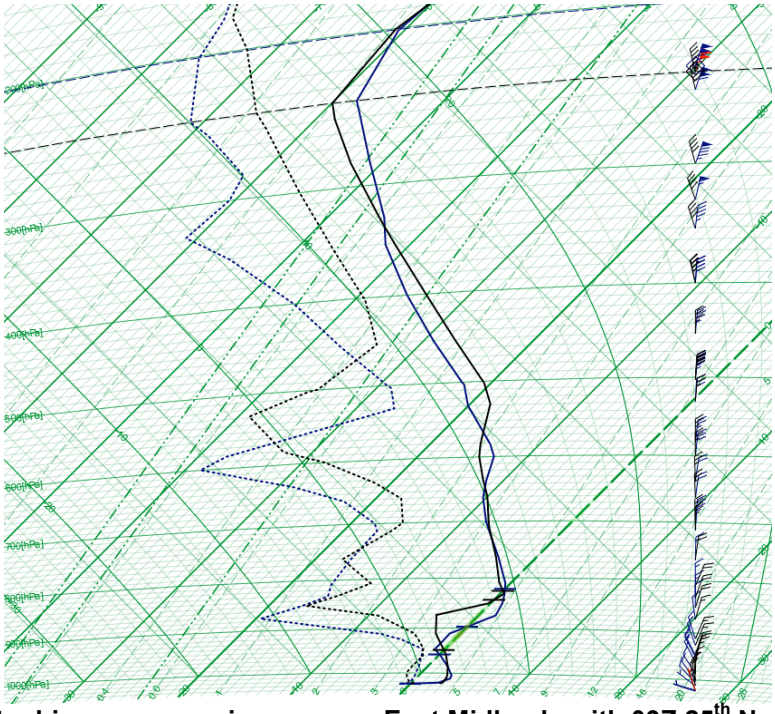


Figure 2 GM tephigram comparison across East Midlands with 09Z 25th November (T+9) in black and 00Z 26th November (T+24) in blue. Boundary Layer depth decreases over night according to GM indicating little impact from short-wave upper trough.

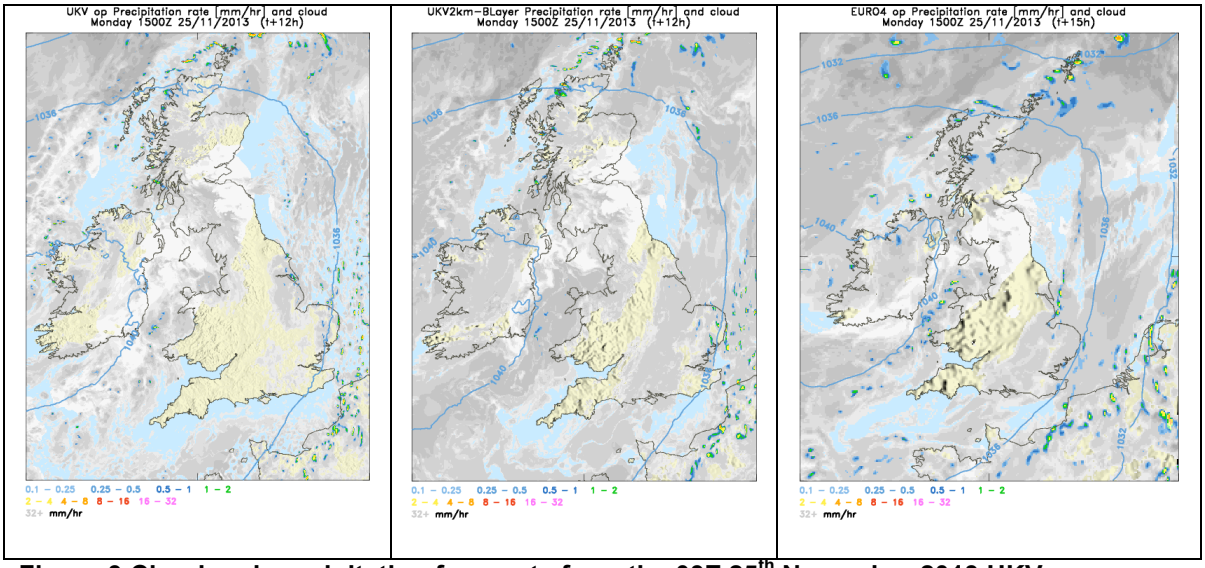


Figure 3 Cloud and precipitation forecasts from the 03Z 25th November 2013 UKV operational model and UKV BLayer models and 00Z 25th November 2013 Euro4 showing major differences across central and southern England.

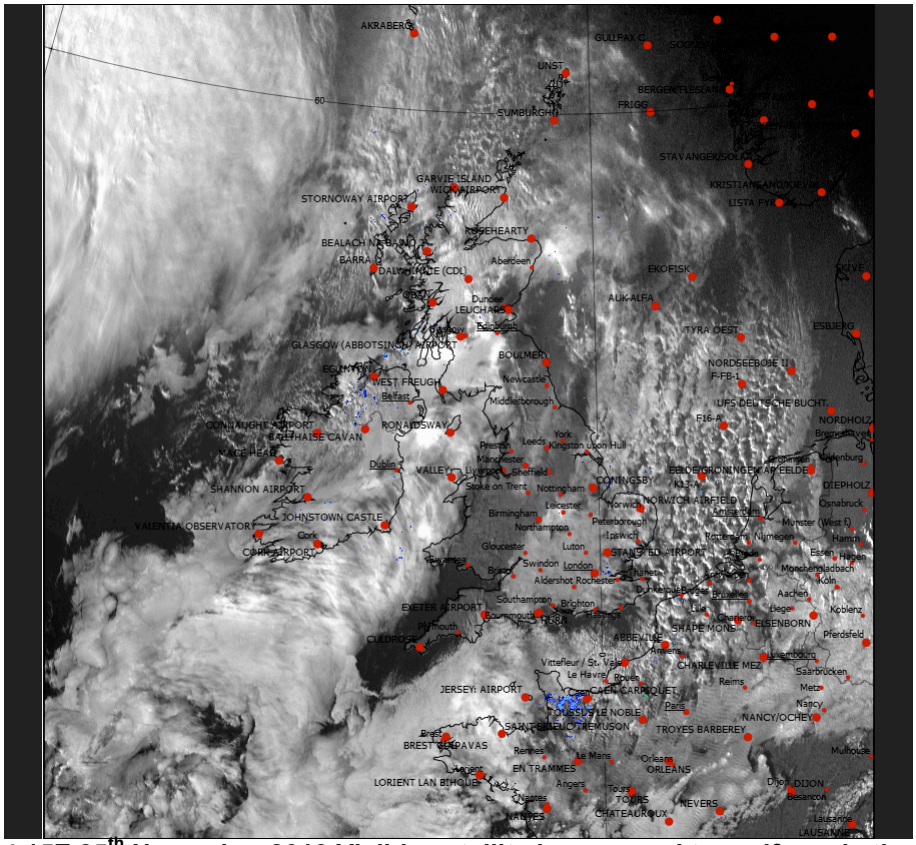


Figure 4 15Z 25th November 2013 Visible satellite image used to verify early time steps of latest model runs.

However, advection techniques suggested that the SW would remain cloudy all day whereas the UKV Blayer indicated that cloud would break in SW later in the day. As the breaking of the cloud in the SW was consistent with

pressure building as the ridge developed, a decision was made to remain with the with the UKV Blayer guidance.

The UKV Blayer and Euro4 models were used as guidance for throughout the short-range forecast period. The models kept low cloud across much of the north and west of the UK but produced clear skies across central England with a clear signal for widespread fog forming across the area during the night (Figure 5), northwest of a line from the Solent to Norwich. This signal was stronger in the UK BLayer model than in the UKV Op and was supported by a strong signal in MOGREPS-UK (Figure 6). With such a strong signal from the models the development of widespread fog and freezing fog was therefore included in the forecast as a significant and widespread risk. However this was not pushed as a certainty and a decision was made not to issue a warning – analysis of model level 6 winds suggested more that that the boundary layer would readily decouple under the ridge conditions leading to frost rather than fog.

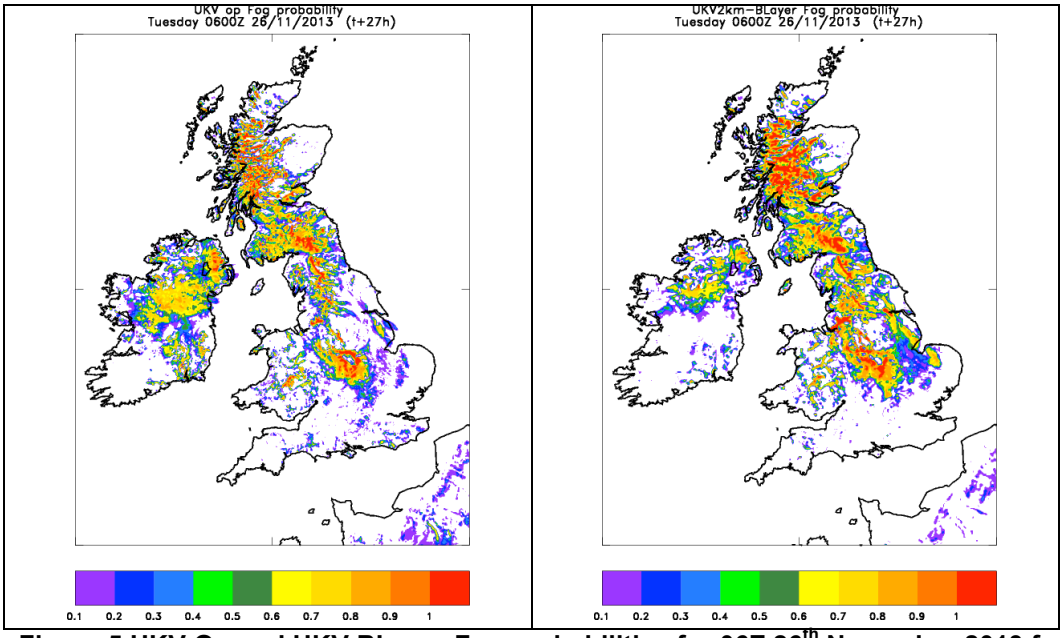


Figure 5 UKV Op and UKV BLayer Fog probabilities for 06Z 26th November 2013 from 03Z 25th November 2013 T+27.

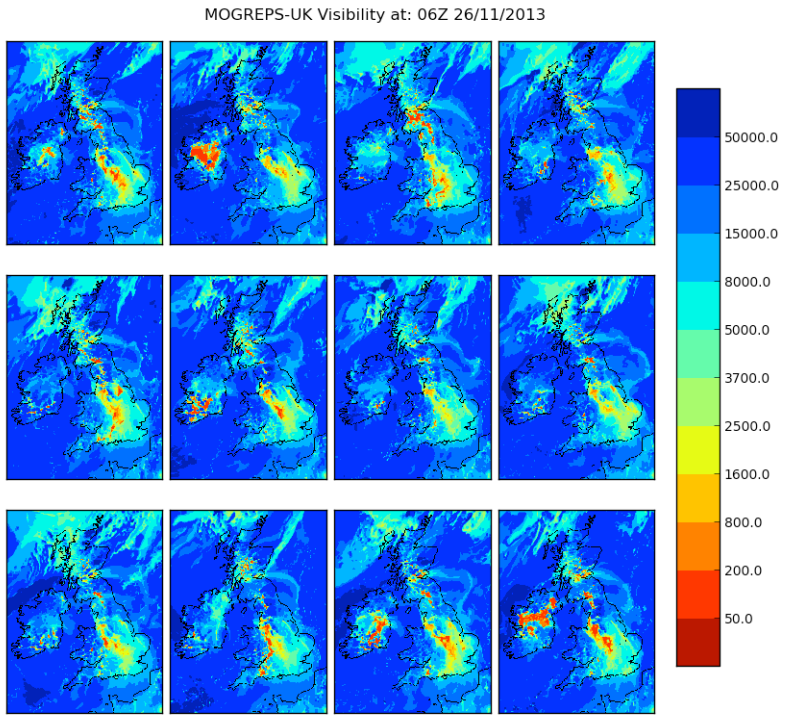


Figure 6 MOGREPS-UK visibility from 03Z 25th November 2013 T+27 showing widespread fog signal across the central England for 06Z 26th November 2013.

Verification:

T+12 to T+18

UKV BLayer and Euro4 models both gave good guidance on cloud cover across the UK. Both models also correctly forecast the breaking of the cloud in the south-west.

UKV Op too clear in the south-east with no representation of the cloud field here.

Good forecast issued for daytime.

T+18 to T+30

All models produced clear skies during the night with widespread fog forming across the Midlands. In reality, whilst some low cloud advected south there was little or no fog and a widespread ground frost.

The 23Z Watnall ascent (Figure 7) indicated that a decoupled boundary layer had developed with little wind and mixing at the surface and moisture being deposited as frost rather than fog. The formation of fog in the models

suggests that the models allowed for too much mixing in the boundary layer. This may be because either the winds at the top of the boundary layer were too strong (generating more mixing than in reality), or else that the inversion in the models and associated decoupling was not strong enough and resulted in too much mixing in the models.

The recognition of this potential for decoupling by the forecasting teams was clearly good judgement against a very strong signal from the models. The decision not to issue a fog warning contrary to the model guidance is further evidence of good judgement. Given the strength of the signal from the models, it is correct that the formation of fog patches was included as a risk.

Poor model guidance across central England for fog/freezing fog development. All models forecast fog rather than frost.

Temperature forecasts for dawn were largely influenced by how well the models had represented the cloud distribution; e.g. Euro4 was too cold in the NW of England and UKV Op was too cold in the SE of England where the models had insufficient cloud.

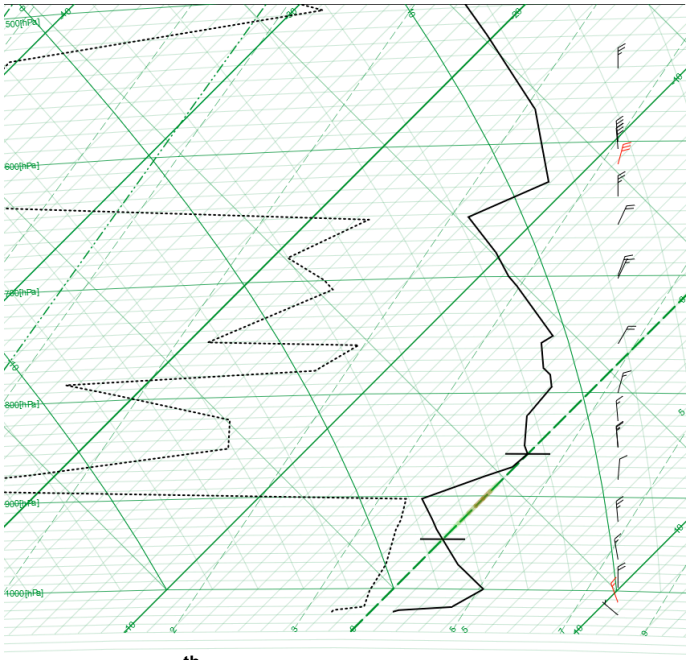
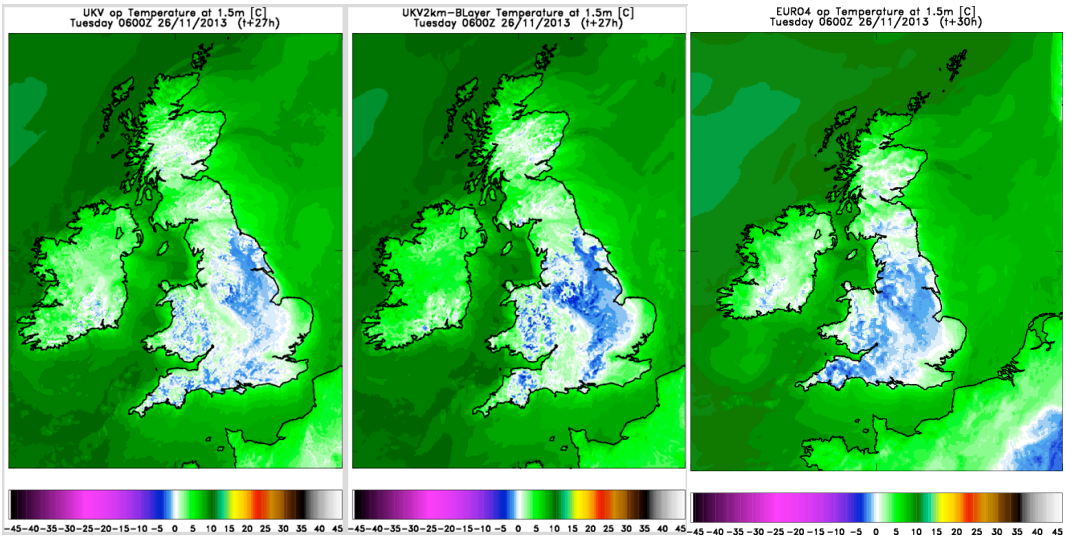


Figure 7 23Z Watnall Ascent 25th November 2013 shows decoupled boundary layer with little wind/mixing at surface, indicates that moisture would have been locked in as frost rather than allowing fog to form.



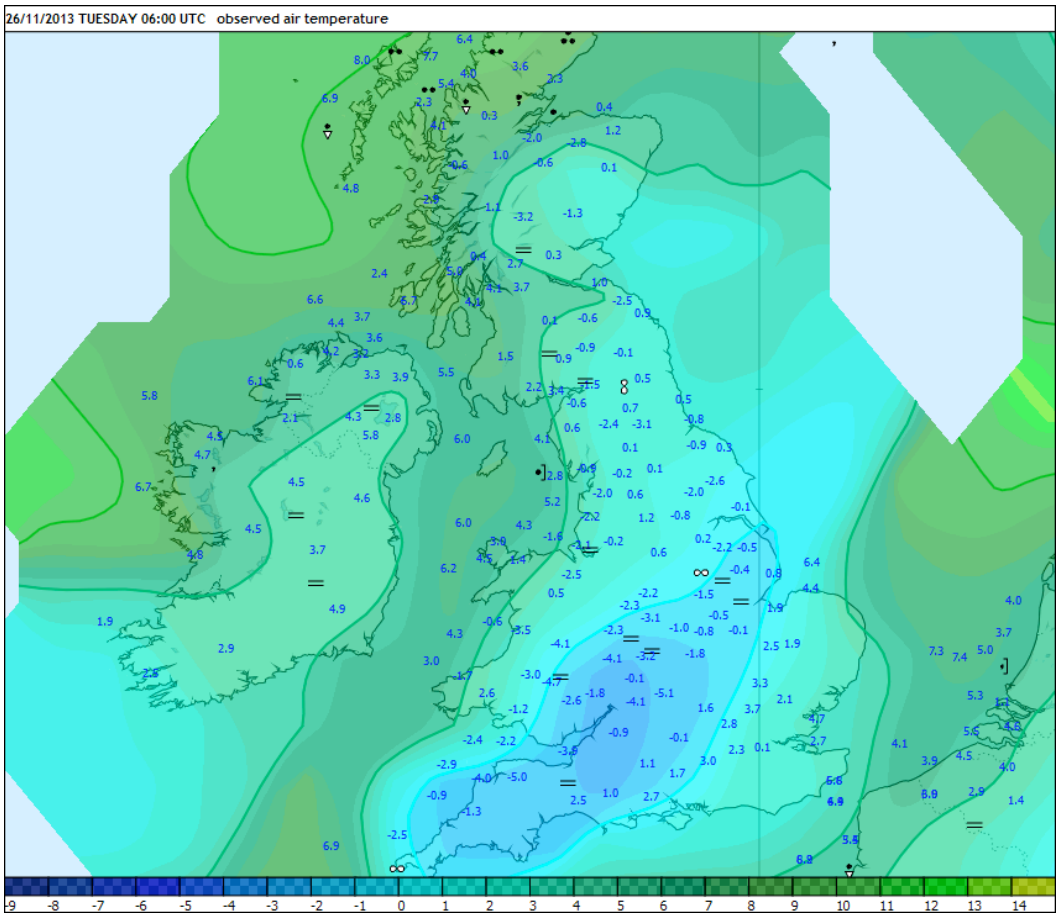


Figure 8 Comparison of 1.5m temperatures from UKV Op (T+27), UKV BLayer (T+27) and Euro4 (T+30) with observed air temperature at 06Z 26th November 2013.

Corrupted Visibilities by Post Processing

Forecast Issued 28th November 2013

Fog across many central and south-eastern areas of the UK is over-forecast by UKV Op, UK Blayer and Euro4. ***Visibility errors are then considerably amplified by the post-processing*** *– many areas' visibilities forecast to reduce to <200m, with some <50 after post-processing.*

Figure 9 shows a comparison between the UKV OP, UKV Blayer, Euro4 and UKV post-processed visibilities from the latest runs available on the 28th November 2013. Figure 10 shows the observed visibilities for verification.

The comparison shows significant errors in visibility for all models – the UKV Op and UKV BLayer are similar for the most part, with reduced visibilities across central and south-east England widely forecast and with 50-100m forecast across Cambridgeshire, north-west England and the Scottish borders. This signal is then amplified considerably once the raw UKV is post processed with <50M visibility widely forecast across central and south-eastern England, north-west England and through the Irish Sea.

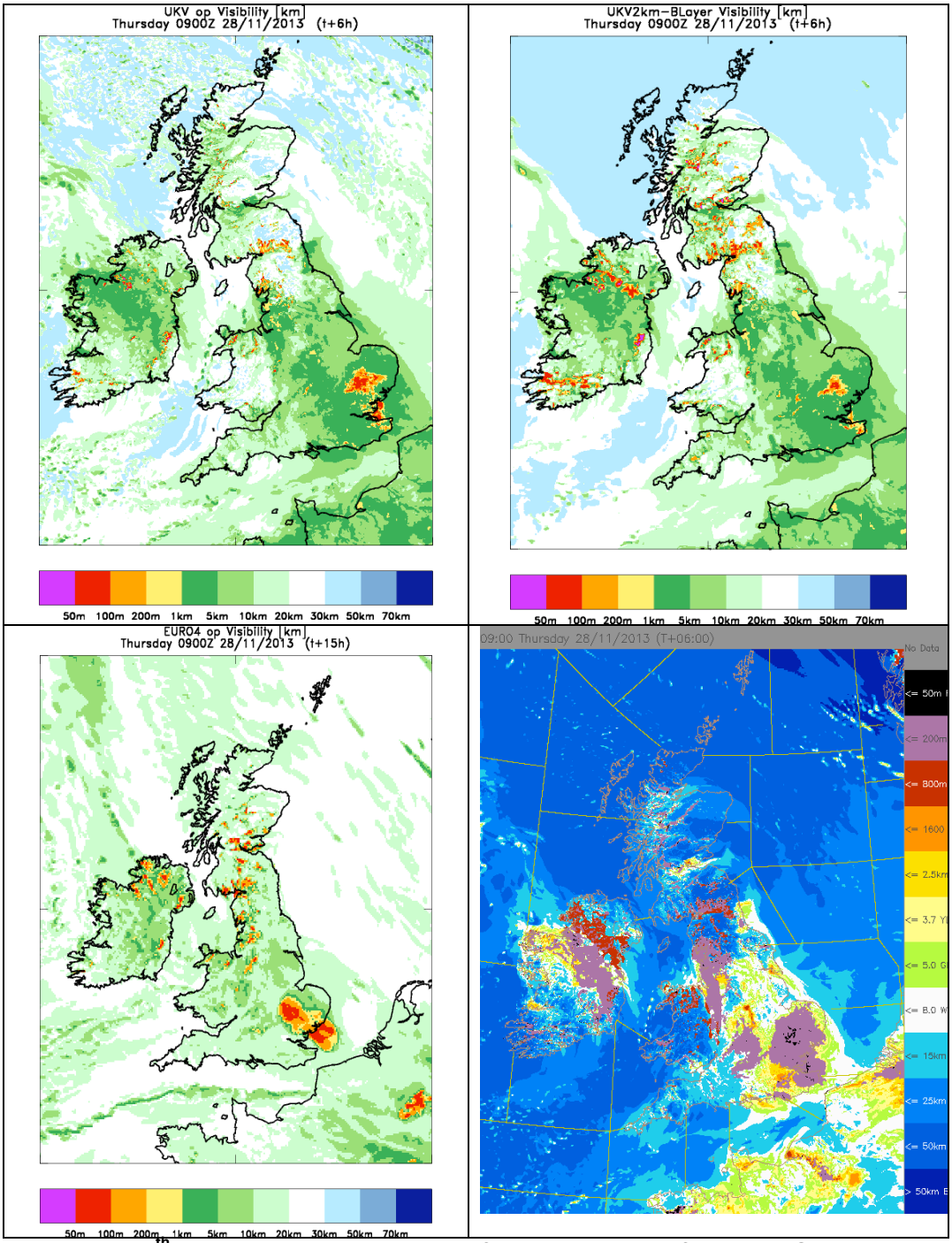


Figure 9 28th November 2013 Visibilities from latest runs of the UKV Op, UKV BLayer, Euro4 and Post-Processed UKV Op.

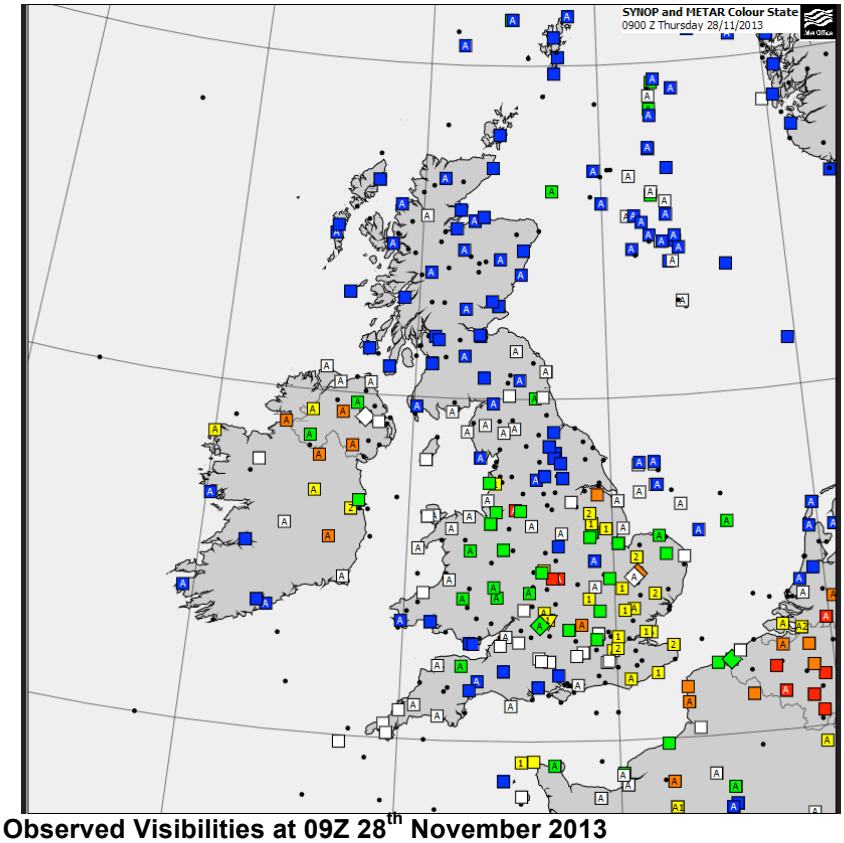


Figure 10 Observed Visibilities at 09Z 28th November 2013

Forecast temperatures too low under ridge conditions and clear skies – decoupling in light wind regimes

UKV and UKV BLayer models both over-forecast temperature falls within the ridge conditions, with the experimental UKV BLayer model indicating larger errors than the operational model. A suspicion developed during the experiment that the UKV BLayer error was linked to the readiness at which teh boundary layer is decoupled in these wind regimes. ***This case illustrates how operational and experimental models can be effectively evaluated in the test bed environment.***

Forecast issued 22nd November 2013

It was observed during the experiment that the UKV and UKV BLayer were both prone to forecasting temperatures that were too low under ridged conditions and clear skies. Figure 11 shows the 1.5m temperatures at T+5 for 08Z 22nd November. Both the UKV Op and UKV Blayer show widespread sub-zero temperatures across Ireland and Scotland. The UKV Op shows temperatures -1 to -3C broadly with some areas of Sotland and Ireland down to ~-5C. The UKV BLayer is much colder with temperatures more widely down to -5C and areas across Glasgow and Edinburgh -10C to -12C.

Figure 12 shows the observed temperatures 08Z 22nd November 2013. Temperatures across the bulk of Ireland are more typically above 0C, with

some central areas between 0 and -1C. There is one location in central Ireland at -5C. Central Scotland temperatures are broadly -2 to -3C, with the coldest temperature of -5.7C confined to the Grampian Mountains. The Glasgow and Edinburgh areas are ~-2 to -3.4C.

Both models forecast temperatures which were in general too low. UKV Op were too low by typically 2-3C but in some areas such as southern Ireland by upto 7C. The UKV Blayer model was more extreme, with temperatures far more widely 5-6C lower and with areas such as Glasgow and Edinburgh upto 12C too low. The UKV Blayer also widely forecasts temperatures too low across western England and Wales.

Clearly the UKV Blayer model was observed in this case to be more prone to over-forecasting overnight minima. This was seen in a number of cases during the course of the experiment. There was a suspicion that the Blayer configuration more readily decoupled the boundary layer so that in marginal situations with light gradient winds an inversion would be created too readily. Boundary layer mixing would therefore be insufficient and temperatures would be forecast as too low. The error would be most marked when light gradient wind (perhaps 10-15Kn) should maintain the mixing and keep temperatures up. In these cases the UKV Op would incorrectly be warmer than the UKV Blayer. In cases of no gradient wind the UKV Blayer configuration and UKV Op would behave similarly with both models generating an accurate inversion.

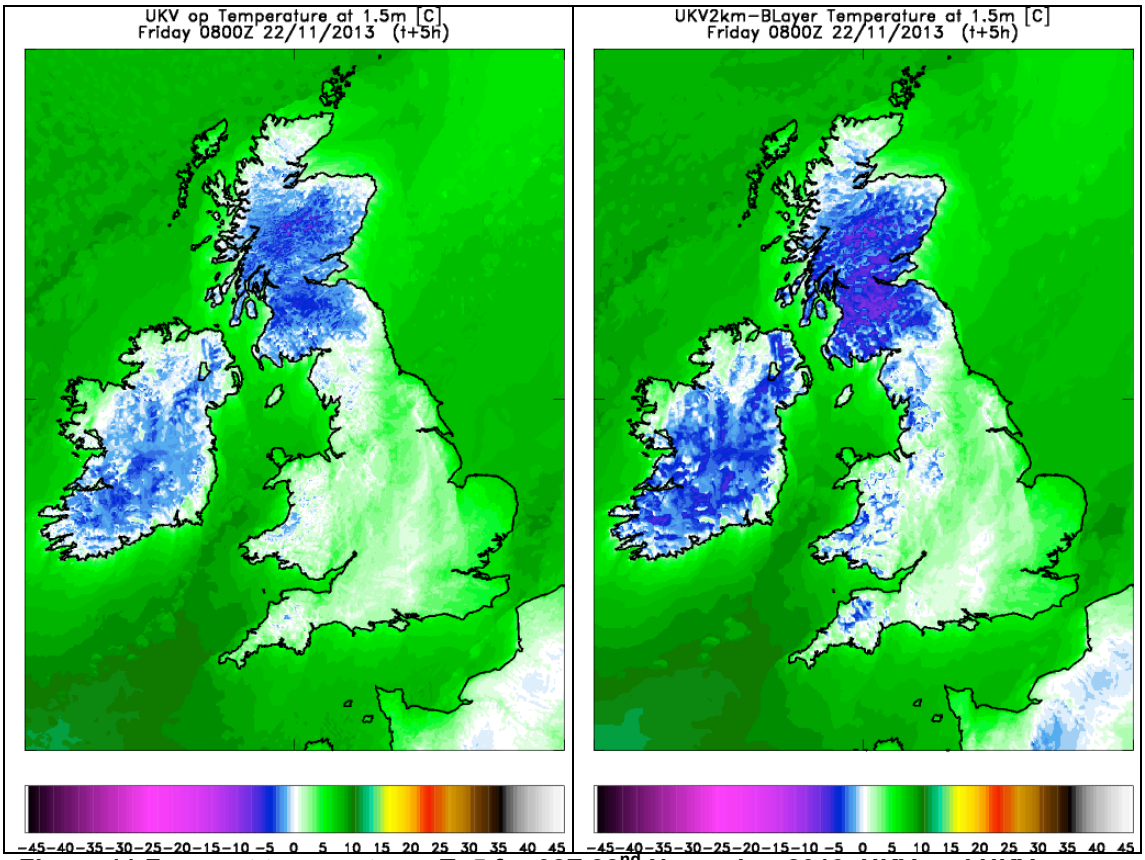


Figure 11 Forecast temperatures T+5 for 08Z 22nd November 2013, UKV and UKV BLayer.

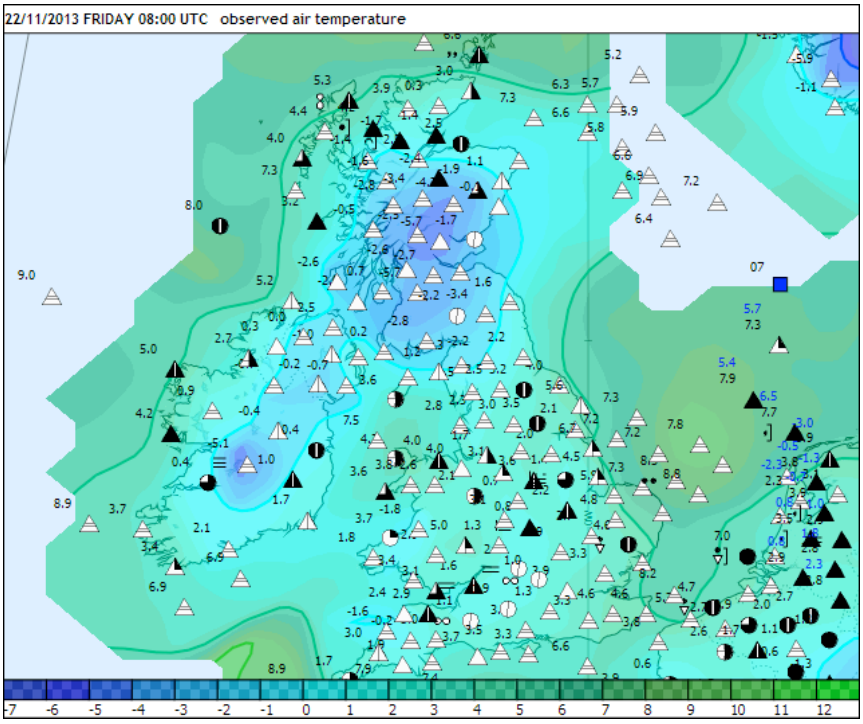


Figure 12 Observed air temperature 08UTC 22nd November 2013.

6. Summary of Findings

Evaluation of Existing and Experimental Met Office Modelling Capability

The following lists summarise evaluations made by the forecasting groups during the experiment. The focus here is on model/products characteristics to be improved rather than the 'business as usual' forecasts of normal skill.

Gp 1 and Gp 2 refer to entries in Group 1 and Group 2 evaluation forms respectively with date dd/mm indicated. ATS refers to a note made specifically by the author.

UKV Op v UKV BLayer Observations

- Much colder signal of min temperatures in BL – 18/11 Gp 1, 21/11 Gp 1, 26/11 Gp 2. Used mid-way UKV-OP UKV-BL 21/11, 26/11 (mid-way to counter known UKV-OP warm bias in min temps).
- More cloud in BL c.f. UK V Op – 20/11 Gp 1, 21/11 Gp 1/ATS, 25/11 Gp 1. UKV breaks cloud too readily – 25/11 Gp 1. BL often verified better for cloud, or solution between UKV-OP and BL – 22/11 Gp 2 verif, 25/11 Gp 1 verif, 26/11 Gp 2 verif. (UKV-Op more likely to break cloud than BL, BL more likely to maintain cloud cover).
- BL and UKV-OP fog risk too large, BL tending to be more so - 21/11 Gp 1, 22/11 Gp 1, 26/11 Gp 2
- BL overnight minima observed to be too low by ~3-4C on clear nights with slight gradient but better than UKV-Op on still nights (threshold of gradient wind ~10-15 KN); suspicion that (a) BL creates a low-level inversion too readily, thereby decoupling the boundary layer too readily leading to lower surface temps or (b) downward eddy flux (mixing) is insufficient. – e.g. 22/11 Gp 2
- Visibilities too pessimistic in UKV especially in Liverpool/Manchester area; due to MURK dataset (too much pollution at source locations).
- UKV BL general gives a better indication of light ('spotty') precip areas (UKV Op tended to overdo this in Sc situations, but be deficient in other regimes). E.g. 27/11 Gp 2.

UKV Op v Euro4 Observations

- E4 generally gave better guidance than UKV-Op during the ridged weather regime dominant during much of the experiment (UKV-Op models the convective overturning in the ridging and breaks-up the cloud too readily; the coarser E4 does not.).

- UKV-Op under does post-front cold air convection, Euro4 better – 29/11 Gp 1
- Euro4 advects fog into sea with suppressed temperatures (5-7C, OSTIA temps 8-10C, Ship 10.2C) – 28/11 Gp 1

Known Model Characteristics Observed

- UKV tendency to break up Sc sheets too readily – 25/11 Gp 1
- UKV warm bias resulting in reduced overnight minima - Gp 1 19/11
- UKV radiation cooling rate too large – too rapid falls of temperatures results in larger temp errors early in the night; not so great later in night so final model error not so large (but suffers from warm bias above) – 21/11 ATS
- UKV/E4 too readily generates low visibilities.
- Euro4 overdoes convection as it organises it onto the grid – e.g. Gp2 20/11 MR.
- Global Model tends not to break up cloud sufficiently and produce too much low low cloud. This sometimes accompanied by excessive 'spotty' ppt.
- Problem with Gust Diagnostic for UKPP/MOGREPS/E4 – 19/11 Gp 2.
- Dry Bias in ECMWF deterministic model – 22/11 Gp 2
- UKV/UKV BL tend to develop fog too quickly (i.e. fog predicted to form too early in the forecast relative to reality) – 28/11 Gp2.
- Simulated Satellite imagery for Global Model appears too moist – this due to excess cirrus/ice cloud aggregation in the GM. (Gp 2 26/11)
- PP Fog signal grossly overdone. (Gp 2 28/11).

Notable Poor Forecasts

- Widespread dense fog incorrectly forecast across central and Northern UK, UKV-OP, UKV-BL, E4, 25/11 for 26/11. All models predict too large fog signal. This reinforced by MOGREPS probs for fog in areas. UKV-PP increases signal. Gp 1/2 25/11, Gp 1/2 26/11 verif.
- Poor handling of fog/visibility across E and SE, 28/11 for 28/11. UKV-Op/BL too pessimistic, UKV-Op especially so, UKV-PP worse still. London Model also. E4 better. UKV-Op shows advection of fog across Thames Estuary. EPS-W MOGREPS-UK shows erroneous output for visibility. UKV-Op/UKV-BL/E4/London, 28/11 Gp 1/2

Identified Process Changes

Improvement in accessibility of Model Characteristics Documentation.

Suggested Product Improvements

GPP

- Interface to keep the selected model run time when changing selected variable instead of defaulting to 'Most Recent'.
- MOGREPS-UK output on GPP postage stamps to include low cloud fraction as an option rather than total cloud only.
- Indication of which model is being displayed in the title of Seamless UM product.
- Request cloud/precip/MSLP charts similar to those on Clive Wilson's pages.
- Request temperature anomalies map (Temp-Climat.).
- Extend seamless UM animations by including GM from T+48.

Simulated Satellite Imagery

- Inclusion of fader option rather than a slider in simulated WV imagery.
- Inclusion of radar/model precipitation with simulated WV slider.
- Alternative greyscale option for simulated WV in cold airmasses to improve contrast
- Apparent parallax error in simulated satellite imagery (e.g. Gp 2 26/11)
- Simulated satellite imagery for North America (requested for and supplied during the Experiment).
- Facility to overlay some fundamental model fields onto imagery (e.g. mslp. 500hPa gph).

Clive Wilson's Model Comparisons

- Cloud representation – not clear on the correct interpretation of the cloud fields. Request notes on correct interpretation. (Example reference - Gp 1 26/11).
- Screen temperature scale shows insufficient detail over the UK (ranges from -45C to +45C)
- UK zoom of Euro4 on same projection as UKV.
- Rachel North's model tephis to contain BL zoom.
- Rachel North's model tephis to include Torshavn (useful during cold air outbreaks).
- Rachel North's tephis – add hi-res tephis?

Forecast Monitoring

- UK Zoom

- North America Zoom
- Analysis Increment plots to return.
- 6 hrly frames in the Model-Model Difference Tracking.
- Investigate logarithmic scale for Different Tracking.
- Investigate slider for model comparisons.
- Add GFS to Forecast Monitoring.

Ensembles

- Larger North Atlantic area (MOGREPS-G) – i.e. western boundary further west
- Tephigrams (MOGREPS-UK)
- EPS-W tick box entry to identify what 'yellow/amber' is showing.
- EPS-W differences in likelihood percentages c.f. NSWWS impact matrix.
- EPS-W thresholds perhaps require review? – in some cases indicated warnings too readily.
- Regime Analysis – consider possibility of showing postage stamps of members when hovering over the number for each day.
- MOGREPS total cloud amount or composite for postage stamps.
- Scales on Previn ECMWF perturbations needs addressing as fields appear washed out.
- Calculation of spread of ECMWF ensemble at T+0 and displayed on same scale as MOGREPS spread at T+0.
- Wind Gust Probabilities charts have different colour schemes depending on whether GM or ECMWF is selected. Useful to have same scales.
- Cyclone Database – UK Zoom.
- Cyclone Database – Consider possibility of clickable dalmation plots.

Suggested New Product Requirements

- Forecast data presented in METAR form for direct comparison with METAR obs.
- Simulated ECMWF Satellite imagery.
- GFS model included in Forecast Monitoring.
- GPP Slider functionality for cloud layers – this to be able to display all cloud layers on a single chart and use a slider to select display of pre-defined cloud at particular heights (i.e. ascend or descend through cloud layers).
- Better representation of cloud in GPP; currently does not reflect optical significance of features such as thin cirrus. (This was discussed on the experiment with Keith Williams and David Walters, who informed us work is currently being done on this subject). Included here as an indicator that interpretation of the cloud cover during the experiment was one of the major issues highlighted in the experiment.

Noted Product inefficiencies/limitations

- Timeliness of EPS-W (MOGREPS-UK) not in at 10:45 from latest 03Z run. Can this be made available earlier?
- SREWS in Visual Weather triple totals for anything other than hourly totals.
- Speed of Visual Weather in Training Rooms 1&2

Noted Products which Decreased Confidence

- Interpretation of low spread around driving model in ensembles when ensemble solutions differ. (Individually low spread indicates high confidence, but different low spreads indicate low confidence). Guidance from research requested.

7. Outcomes

Process

- **Developed, evaluated and refined a process for experiments** of this type to be done in the future.

Communication

- **Increase in integration and communication** between operational meteorologists and research scientists both as the larger units and also on more personal levels.

Models & Products

- Evaluation of current and experimental models by collaborating forecasters and scientists gives **useful and unique insight to Science** for the future development of our modelling capability.
- Forecasters made aware of new or existing products enabling more **rapid and widespread take-up** and with forecasters directly being able to affect the way products are developed further.
- Scientists reported being made aware of new details in their own work as a result of the experiment thereby providing **different approaches** from which to progress their own research.

- A multitude of existing **product improvements** were suggested that will enhance the clarity, interpretation and use of the products. This will increase the product usefulness and forecasting efficiency. Some changes were implemented within weeks of the end of the experiment.
- Many **new product ideas** were suggested that address key deficiencies in the information available to the forecaster thereby enhancing the model guidance.

Knowledge & Development

- The Experiment highlighted the **critical importance of the model characteristics** so that they could be accounted for in the consideration, interpretation and selection of different models when producing a final forecast.
- Judgement calls made during the forecasts were an important aspect under consideration. Discussion both during and after the group forecasting and verification activities provided an **exchange of knowledge and experience which were unique to this environment**.
- **Outstanding personal learning and development** opportunity for forecasters and scientists alike with exposure to new and wide-ranging products, techniques, knowledge and experiences.
- Identified areas of confusion and ambiguity in the interpretation of model products, thereby identifying either **training needs** or **clarification** of product displays.
- Identified the most **useful, informative and efficient products and techniques**. This could be developed to refine the recommended process for operational forecasters.

8. Benefits

- Encourages interaction and increases communication between Scientists and Forecasters thereby **improving the strategic direction** of the Met Office.
- Provides valuable information on skill and efficiency of forecasting and scientific techniques so as to examine ways of **improving forecast accuracy and lead-time** and help **guide research and development strategies**.
- Early exposure of forecasters to new and developing scientific products helps to **increase usefulness of products** by helping to guide product development.
- Scientists gain an understanding of the challenges of operational forecasting and an insight into how to **improve future forecasting** products and techniques.

- Operational Forecasters gain a greater understanding of the scientific principles involved in the numerical models and an insight into making **more skilful interpretations** of numerical models with a greater knowledge of model characteristics.
- Newly acquired knowledge by forecasters and scientists is communicated back to colleagues in own areas so that **experience is distributed** beyond those attending the experiment.
- **PWS** involvement ensures **customers' evolving needs are represented** in the development of new products, services and messaging. PWS are made aware of the limitations of current developing science and forecasting techniques aiding in its **communication to customers**.

9. Recommendations

- **Model characteristics should be well documented, remain current and be readily accessible to all forecasters and scientists**. The experiment highlighted how a knowledge of current model characteristics were critical to the forecasting process.
- **Make greater use of model upgrades, experimental models and product developments** where possible for future experiments to give the opportunity for their evaluation in the correct environment. Evaluation by on-duty operational forecasters is inadequate due to lack of time on shift.
- **Instances of poor forecasts highlighted during this and future experiments should be recommended to research groups for further investigation** or merged with their generic problem in the NWP problem group database.
- **Product improvements or new products to be disseminated** to relevant research groups for their consideration.
- **Develop a recommended guidance unit forecasting process** to refine existing practices based on the experiment outcomes. (This could be modified in future experiments to evolve as new products and techniques appear or as old ones become superseded).
- **Involve a representative from the products or business teams in future experiments** so that not only can new products may be identified early but also so that current forecasting limitations are recognised from a business viewpoint.
- **Improve performance of Visual Weather in Training Rooms 1&2**. Speed of Visual Weather in the Experiment rooms needs addressing for future experiments.

10. Lessons Learned

Verification

By including the medium range in the experiment the assessments became too focussed on the requirement to generate the forecast for all the times from T+12 to T+144. Although this allowed for wide-ranging model evaluations and generated meaningful and fruitful discussions, the time for verification of the models was reduced to below a level that was optimum. In trying to consolidate the information from the forecast exercises the author found that some of the most valuable information from this experiment was in the pre-forecast assessment of the current model runs (i.e. the verification of the early timesteps of the latest run) which evolved out of the experimental forecasting process as the experiment continued.

In future experiments it is suggested that the remit be shortened so as to focus on a much narrower timeframe – either T+12 to T+48 to focus on the short-range high resolution models or else perhaps T+48 to T+144 to focus on the long-range and ensembles. This will allow for a more concise experiment targeted on more specific modelling challenges and allow for a more significant period of verification of the models.

Cloud Verification

Verification of cloud during the experiment was a difficult exercise. The production of new products to aid in this would be beneficial. For example and overlay of 6/8 cloud cover on Satellite Imagery could be produced in Visual Weather.

Daily Weather Summary

A concise weather summary should be written at the end of each day to help in the verification aspects both during and after the experiment.

Evaluation Questions

The questions designed for the evaluation should become more focussed so as to steer the answers along more useful and productive lines. A use of subjective numerical ratings should also be used more – e.g. “model X was better/same/worse than model Y for this forecast for temp/clouds/vis/precip”. These can then be built into statistics of the performance of an experimental or new product/model relative to an existing one in the post-event for analysis.

Evidence

It should be stressed to the participants that as much supporting evidence should be included as possible in the documentation with model fields. Although this sounds obvious there were a number of occasions in the experimental documents in which this was not included sufficiently.

Warnings

The warnings matrix was included in the evaluation as the experiment progressed as a concise way to judge warnings success. This was a successful inclusion in the process and it is suggested that the 'warnings cross' is indicated whether a warning is issued or not along with a record as to why the decision was taken.

IT

The speed of IT in the college training rooms was very slow at times. Visual Weather in particular was largely unusable so much of the analysis was performed via webpages. The performance of Visual Weather seems to be due to both the College data servers (which have less resources than those in Ops Centre) and the workstations (which are comparatively old compared to Ops Centres).

Ideally we would replicate Ops Centre as closely as possible in terms of IT infrastructure.

Layout of IT for the Forecasting Groups

The layout of the workstations for the two forecasting groups worked well. The main computer displaying to the big screen for the group to see was effective. The workstations to the side became the base for filling out the assessment forms. It is worth experimenting with this layout further and moving the two side workstations closer to the main station – this because at times (and depending on group dynamics) the groups could fragment between the two workstations. Bringing the workstations closer together would help eliviate this.

A further issue is that the wealth of data available requires multiple screens. Perhaps a total of six screens close together should be considered.

11. Summary

The Autumn Experiment was the first forecasting experiment carried out in the Met Office in which Research Scientists, Operational Meteorologists and Public Weather Service Advisors worked together in a simulated operational environment in order to critically assess the Met Office's forecasting capability. The experiment was inspired by the 'weather test beds' which have developed in the last decade in other meteorological establishments around the world but which had never been investigated here.

The Autumn Experiment was therefore primarily a feasibility and suitability study in conducting this type of exercise at the Met Office, and also to establish a process from which future experiments could be developed. In these aspects the experiment was successful: the test bed environment was shown to be entirely suited to the Met Office's forecasting infrastructure and a blueprint for the experimental process was developed for future experiments. The experiment was successful in showing how the environment could be used to productively evaluate developmental models (in this case an experimental version of the UKV), and also in delivering recommendations on existing and future forecasting products.

The experiment also provided an insight into the less tangible benefits of this exercise. Above all the experiment facilitated an increase in communication across Science, Forecasting, PWS, and Training, both before, during and after the event. If this initiative is carried forward with further experiments then it can only help to build new networks across all participating areas of the Met Office. Further, with the combined participation of Science and Forecasting this will aid the future strategic direction of the Met Office as models and products have operational meteorologist involvement from an early stage.

One aspect which became apparent as the experiment progressed was how important a knowledge of model characteristics were to an effective forecasting process. The model characteristics information was critical for the assessment and evaluation of the different signals from different models and therefore in the making of an effective judgement for the final forecast guidance. The implications are that not only must operational meteorologists be aware as much as possible on the latest and developing model characteristics but that these characteristics must be kept current, well documented, and accessible to all.

All participants of the experiment have spoken very highly of the experience and many have voiced a desire to return on future experiments; the experiment provided a fantastic learning and development opportunity for all involved.

A significant number of lessons have also been learned as a result of carrying out the experiment which will help plan and develop the idea in future experiments.

Perhaps the most important of these is that the balance of real-time assessment activities and verification activities was not right this time with too much time being spent on real-time forecasting activities and not enough time on forecast verification. The real time forecasting exercises are valuable in promoting discussion about which model is being most effective in respresenting the situation well and also in stimulating conversation as to judgement calls (and in this aspect this is where the model caharcteistcis is so important). The forecasting exercises are also extremely valuable in evaluating how efficient model products are and in recommending efficiency and clarity improvements. However it is only in the verification exercises can the participants truly experience which model or model product was more effective and in this way provide one of the best learning opportunities for all concerned.

Another aspect which will be changed in future experiments will be on the forecast ranges involved. For reasons of the pilot study discussed earlier, the full forecast ranges of the short- and medium-range were considered. However, in the event this was too much to cover when also attempting verify forecasts for both ranges. Future experiments will therefore focus on either the short- or the medium-range forecasts. This will not only provide more time for the forecasting activity but also have the added benefit of releasing the extra time necessary for the verification activities as discussed.

In this experiment a number of new model products were used along with an experimental version of the UKV and an experimental London model. The experimental UKV (with a trial boundary layer scheme) in particular was used widely and provided a means to compare an operational model with a model that was in development. In the weather regimes of the experiment this allowed the participants to observe performance differences in aspects such as maximum and minimum temperatures, fog formation and cloud distribution. On this occasion the use of the London model was limited because of the persistent weather regime but this model too would have been evaluated in a similar way given appropriate conditions.

It is these experimental aspects that the next forecasting experiment would seek to grow: model upgrades, parallel suites, experimental models and schemes etc and experimental products would all be consistent with the vision of the experiment.

The Autumn Experiment has been a successful investigation into the use of the test bed initiative at the Met Office and given the resources and opportunity could grow into a core activity for Science and Forecasting.

12. Acknowledgements

The author is indebted to Steve Willington who had represented the Met Office at the Hazardous Weather Test Bed (HWT) in Oklahoma and returned to request funding from PWS so that the Met Office could try a similar initiative. The author is also indebted to Steve for giving me the freedom to plan and organise the experiment in all aspects.

The author also wishes to thank the participants who volunteered for this first Forecasting Experiment; for engaging with the spirit of the experiment, for persevering through the culture shock of the first few days and in helping to develop the experiment into the success it proved to be.

13. References

Ralph F.M *et al* 2013, The Emergence of Weather-Related Test Beds Linking Research and Forecasting Operations, Bull. Amer. Meteor. Soc., 94, 1187–1211

14. Appendix

Forecast and Evaluation Templates

Forecast Period T+12 to T+48

Forecast Group X

Issue DDHH00

1. Latest Model critique against current observations (00z-12z)

Brief assessment of any chosen models in the time preceding the start of the forecast period

2. Forecast Guidance & Evidence

i) Forecast

With each forecast statement state in brackets the evidence and judgement call used. E.g. Widespread harsh frost across southern England (UKV, Euro4 considered too warm because…)

Day 0 day 12z-18z (T+12 to T+18)

- **Weather -**
- **Winds -**
- **Temp -**

Day 0 night 18z-06z (T+18 to T+30)

- **Weather -**
- **Winds -**
- **Temp -**

Day 1 day 06z-18z (T+30 to T+42)

- **Weather -**
- **Winds -**
- **Temp -**

Day 1 day 18z-00z (T+42 to T+48)

- **Weather -**

- **Winds -**
- **Temp -**

ii) Uncertainties
(Discussion of Timing/depth of low/intensity of precipitation uncertainties)

2. Messaging and Warnings

(Main messages and warnings if appropriate)
i) Messaging (state evidence and judgement call)

ii) Warnings (if appropriate) (state evidence and judgement call).
State where the tick would be placed in the matrix if issuing a warning.

Likelihood	**High**				
	Med				
	Low				
	Very Low				
		Very Low	**Low**	**Med**	**High**
Impact					

3. Model/Product/Technique Evaluation

i) Evaluate usefulness of any models/products/techniques used. Were any models/products more useful than others and why? Can the presentation of any products be improved?

ii) Did any models/products/ increase/decrease forecast confidence and awareness? Give reasons.

iii) List any products which made the forecasting process more or less efficient and state reasons. If appropriate, indicate how products could be improved to make them more efficient.

iv) Is there any information which was unavailable to you which may have made a difference? Can a new product requirement be identified in this instance?

v) State any known model problems or biases which have been observed. (e.g. Euro4 warm bias in clear night).

vi) Observations made of differences between UKV and UKV BLayer. Include images as evidence.

Forecast Period T+60 to T+144

Forecast Group X

Issue DDHH00

1. Model critique against current conditions

Brief assessment of any chosen models in the time preceding the forecast period

2. Forecast Guidance & Evidence

i) Forecast

With each forecast statement state in brackets the evidence and judgement call used. E.g. Widespread harsh frost across southern England (Euro4, GM considered too warm because…) .

Be more general in description of weather/winds/temp.

Day 2 day 00z-00z (T+48 to T+72)

- **Synoptics -**
- **Weather -**
- **Winds -**
- **Temp -**

Day 3 day 00z-00z (T+72 to T+96)

- **Synoptics -**
- **Weather -**
- **Winds -**
- **Temp -**

Day 4 day 00z-00z (T+96 to T+120)

- **Synoptics -**
- **Weather -**
- **Winds -**
- **Temp -**

Day 5 day 00z-00z (T+120 to T+144)

- **Synoptics -**
- **Weather -**
- **Winds -**
- **Temp -**

ii) Uncertainties
(Discussion of Timing/depth of low/intensity of precipitation uncertainties)

2. Messaging and Warnings

(Main messages and warnings if appropriate)
i) Messaging (state evidence and judgement call)

ii) Warnings (if appropriate) (state evidence and judgement call).
State where the tick would be placed in the matrix if issuing a warning.

Likelihood	**High**				
	Med				
	Low				
	Very Low				
		Very Low	**Low**	**Med**	**High**
Impact					

3. Model/Product/Technique Evaluation

i) Evaluate usefulness of any models/products/techniques used. Were any models/products more useful than others and why? Can the presentation of any products be improved?

ii) Did any models/products/ increase/decrease forecast confidence and awareness? Give reasons.

iii) List any products which made the forecasting process more or less efficient and state reasons. If appropriate, indicate how products could be improved to make them more efficient.

iv) Is there any information which was unavailable to you which may have made a difference? Can a new product requirement be identified in this instance?

v) State any known model problems or biases which have been observed. (e.g. Euro4 warm bias in clear night).

Forecast Scoring Form - Forecast Period T+12 to T+48

Forecast Group X Scoring Forecast Group Y

Forecast Issued DDHH00
Forecast Scored DDHH00

Forecast Verification Day 0 12z-18z (T+12 to T+18)

Note how the forecast scored for each parameter. Discuss whether the model or product used gave good guidance. Identify if possible what model/product would have given better guidance.

Weather

Winds

Temp

Forecast Verification Day 0 18z-06z (T+18 to T+30)

Note how the forecast scored for each parameter. Discuss whether the model or product used gave good guidance. Identify if possible what model/product would have given better guidance.

Weather

Winds

Temp

Forecast Verification Day 1 06z-18z (T+30 to T+42)

Note how the forecast scored for each parameter. Discuss whether the model or product used gave good guidance. Identify if possible what model/product would have given better guidance.

Weather

Winds

Temp

Forecast Verification Day 1 day 18z-00z (T+42 to T+48)

Note how the forecast scored for each parameter. Discuss whether the model or product used gave good guidance. Identify if possible what model/product would have given better guidance.

Weather

Winds

Temp

Summary

Summarise any significant model performance issues noted above (good and bad).

Was anything missed in any forecast?

15. Testimonials

I feel that I learnt so much during two weeks of just looking at the weather and talking about it. I learnt about some fabulous tools being developed by scientists that have already proven invaluable to me on shift. The non-operational environment of the AFE took away a great deal of stress and pressure and got everyone really engaged and enthusiastic. Having taken part, I would now not allow anyone to underestimate the value of taking a range of forecasters and scientists from all corners of the office, from Chief level down to new entrants, and putting them in a room together for a week or two. It built confidence, relationships and knowledge and I can't wait to see if we will get the chance to do it all again soon!

Forecaster

I have been forecasting for nearly 25 years in a number of different disciplines, including civil and military aviation, PWS, commercial and media forecasting.

Day to day I normally discuss the weather just with Operations Centre staff. The Autumn Experiment gave me an opportunity to also be involved with the very people who helped create the models and applications I used day by day. I could therefore spend time actively scrutinising and discussing model performance with people both within and related to the forecasting process. This has allowed me to develop a deeper understanding of both model strengths and weaknesses and how models can be practically applied. It also demonstrated how a multitude of communication flows could be harnessed which can better inform and direct how the forecasting process develops from now on.

The Autumn Experiment showed me a vision of the future of the Met Office as one where the various individual functions come together in a complementary fashion. This will help reinforce and inform the development of the components of the Met Office, and the overall organisational strategy for the considerable benefit to all stakeholders.

Forecaster

I found the Autumn Experiment to be an exciting, and totally worthwhile, project to be involved with from a personal point of view.
It reconnected me with the reason why we're here doing the job and brought me back into the 'real world' in terms of how my job could contribute to the wider Met Office direction. The experience will eventually feed through, I hope, to allow our team to be less reactionary in terms of model assessment and to be able to think more in terms of the weather.

Although it was hard work at the time, I am committed to ensure the new ideas and fresh perspective it has brought can be applied to my work in the near future.

Scientist

A hugely enjoyable and productive demonstration of what can be achieved when research scientists and forecasters get together. It has been very useful in informing the future direction of my work and I left the experiment with a greatly expanded knowledge of how the forecasting process works and how what I do as a scientist can feed into it.

Scientist

Participating in this experiment gave me a much better understanding for how forecasters use and interpret both raw and post-processed NWP forecasts, and in particular ensemble forecasts. I am now hoping to take this improved understanding back to my day job and apply it to new forecast products I create.

Scientist

I thoroughly enjoyed the Autumn Forecasting Experiment. It was a fantastic personal development opportunity to spend two weeks working alongside other scientists, operational forecasters, and PWS advisors, and experience the challenges involved in the real-time forecasting environment. In my role I often develop products for forecasters and I found it invaluable to see how these and other tools are used, and how existing and new products could be improved and developed to assist in the forecasting process. I sincerely hope that the AFE is continued for years to come.

Scientist

Very enjoyable and very interesting. A great opportunity for forecasters and scientists to work together and gain a better understanding of each other's work. I learnt a great deal from being able to discuss various topics with colleagues. I'm sure the workshop will help shape new forecast products.

PWS Advisor